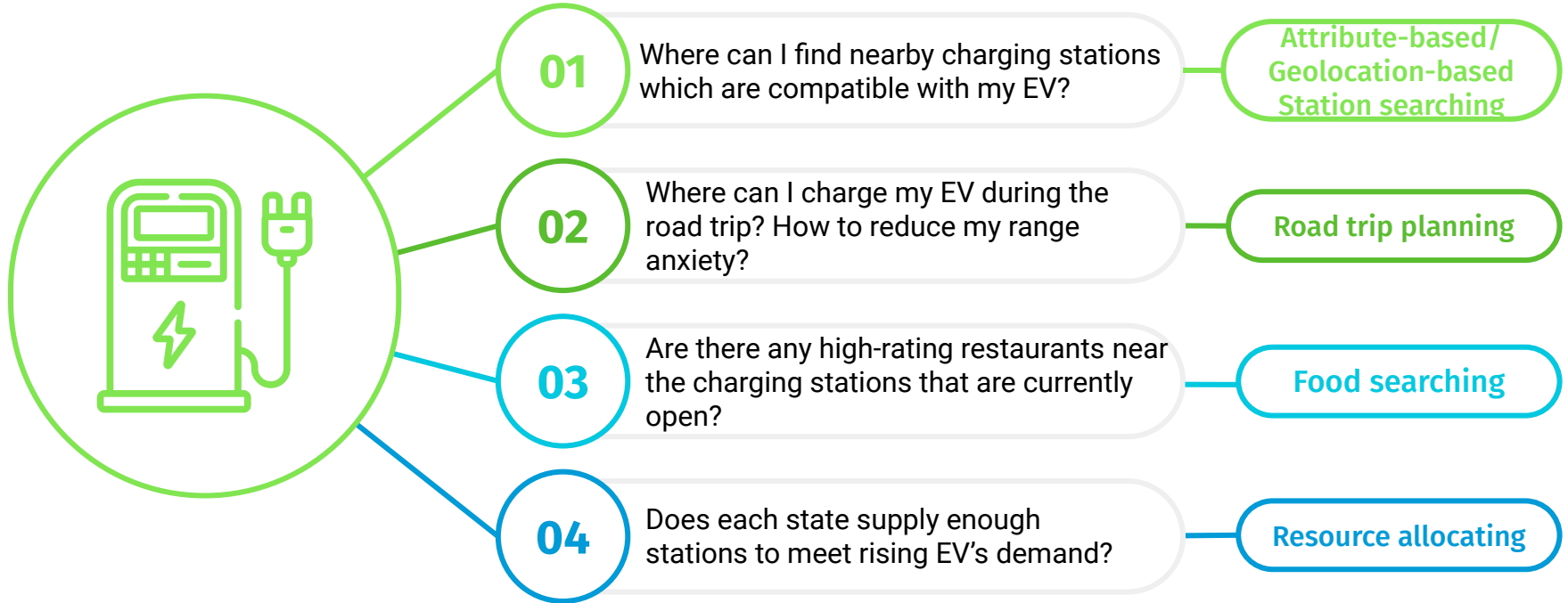


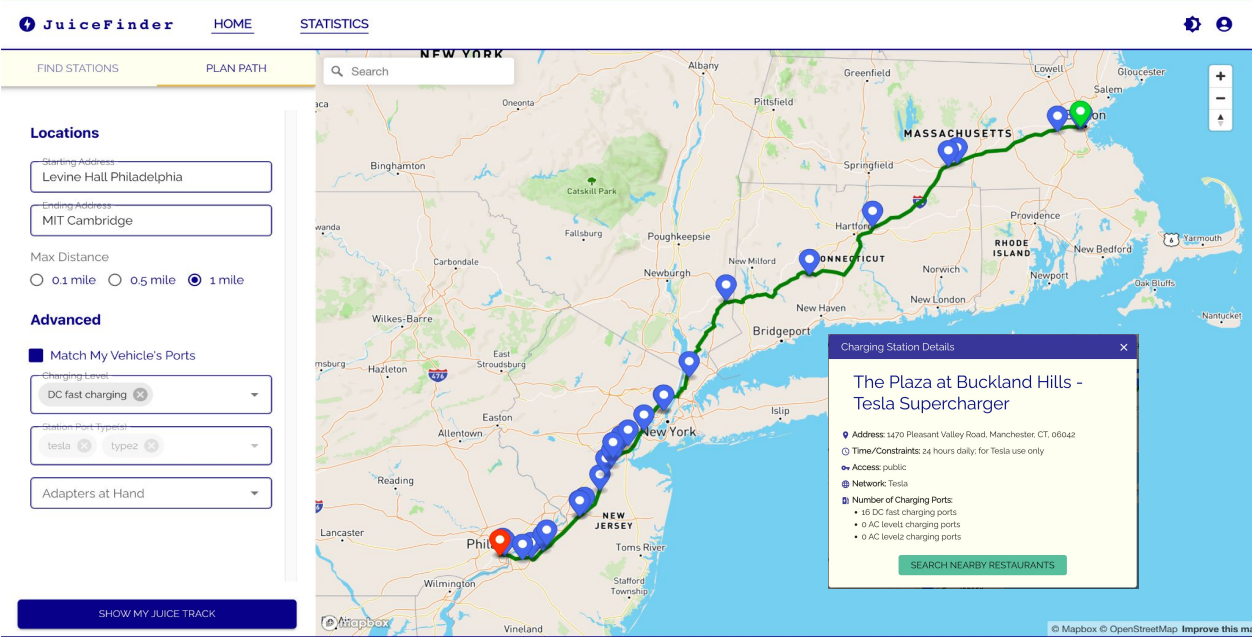
JuiceFinder

Chuan Li, Wei-Shen Lee, Zhiyang Lin, Yicheng Shen

Basic Problems & Goals

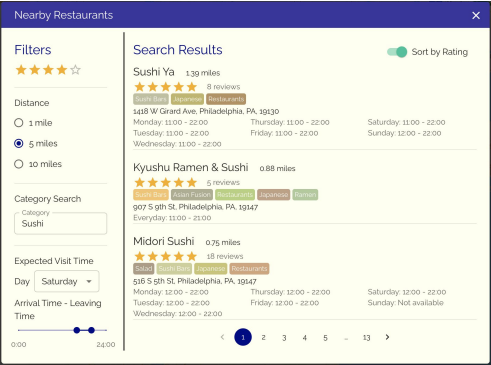
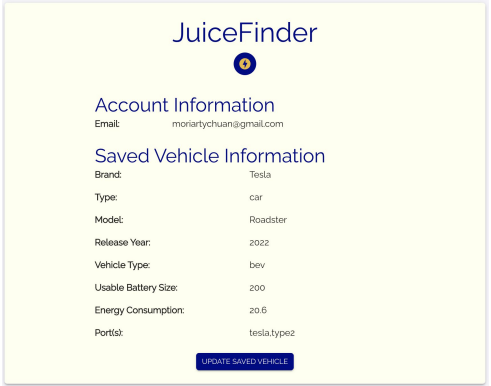


Preview



Copyright © 2023 JuiceFinder

Contributors: (CL) (WL) (YS) (ZL)



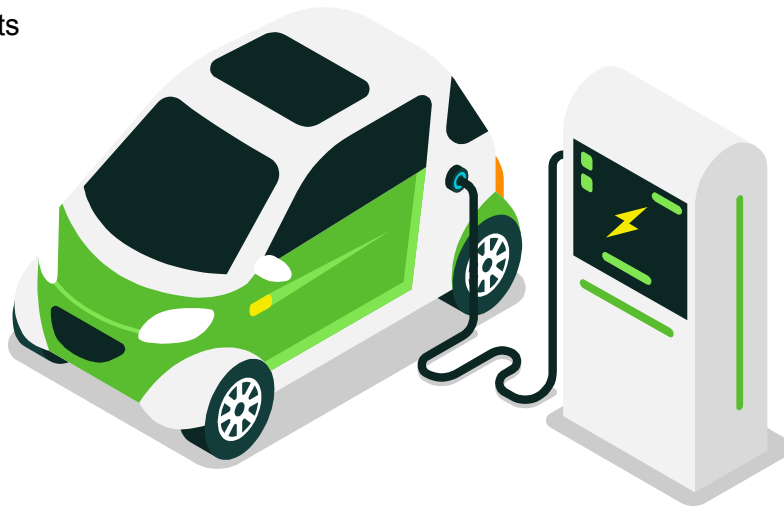
Preview



Datasets

Open-EV-dataset(JSON)

293 objects



Yelp Dataset(JSON)

150,346 rows

US **A**lternative **F**ueling **S**tations(csv)

70,034 rows

Data Preprocess and Normalization

Construct geometric point value using (longitude, latitude) coordinates

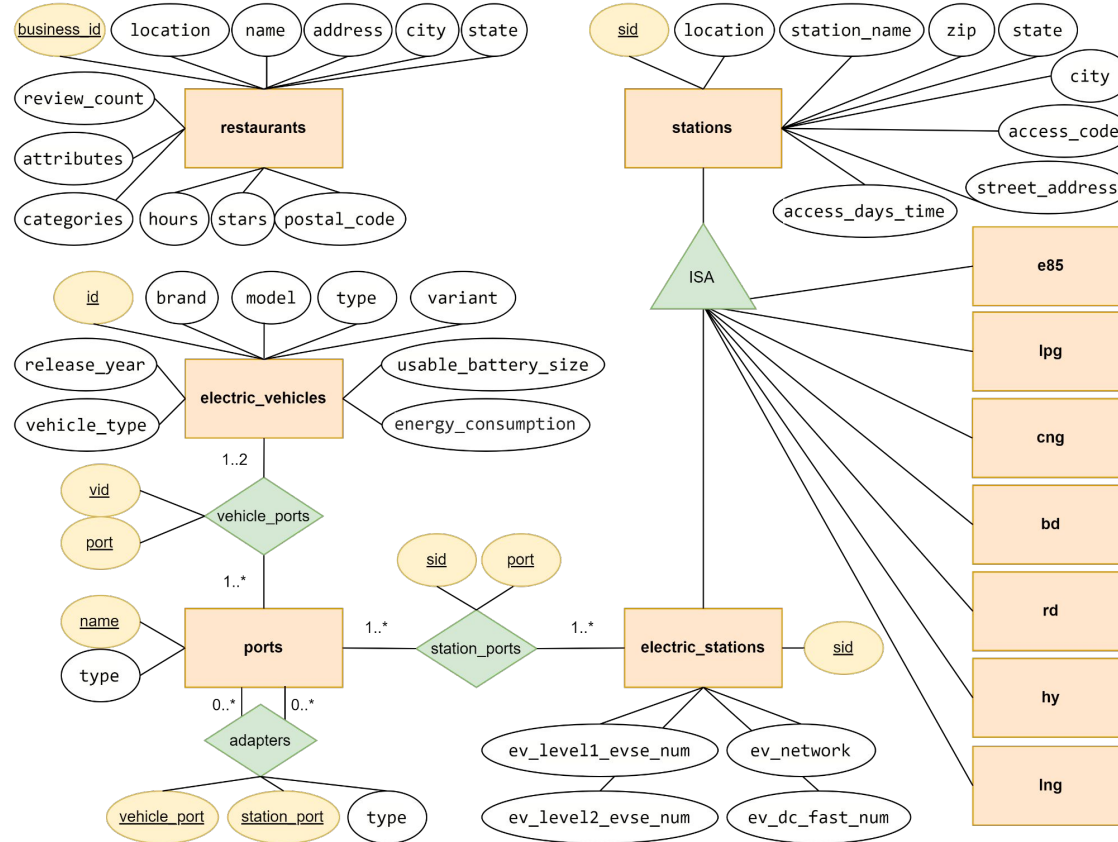
```
insert_data_sql = '''
INSERT INTO Restaurants (business_id, name, address, city, state, postal_code, location, stars, review_count, attributes, categories, hours)
VALUES (%s, %s, %s, %s, %s, %s, ST_SRID(ST_GeomFromText(%s), 4326), %s, %s, %s, %s, %s)
'''
```

Port name unification

AFS	OpenEV	Naming	Compatible with adapter(vehicle to station)
J1772	Type1	type1	tesla_to_type1, type2_to_type1
NA	Type2	type2	tesla_to_type2, type1_to_type2
NEMA515(level 1, slower)	--	nema515	type1_to_nema515,type2_to_nema515,tesla_to_nema515
NEMA520(level 1, slower)	--	nema520	type1_to_nema520,type2_to_nema520,tesla_to_nema520
NEMA1450(level2)	--	nema1450	type1_to_nema1450,type2_to_nema1450,tesla_to_nema1450
AFS	OpenEV	Naming	Compatible with adapter(vehicle to station)
J1772COMBO	ccs	ccs	tesla_to_ccs
CHADEMO	chademo	chademo	tesla_to_chademo
TESLA	tesla_suc	tesla	--
--	tesla_ccs	explode: tesla & ccs	

Every relation is in BCNF because each attribute depends on the primary key.

Schema Design and ER Diagram



Demo

Optimization

Built-In Distance Calculator

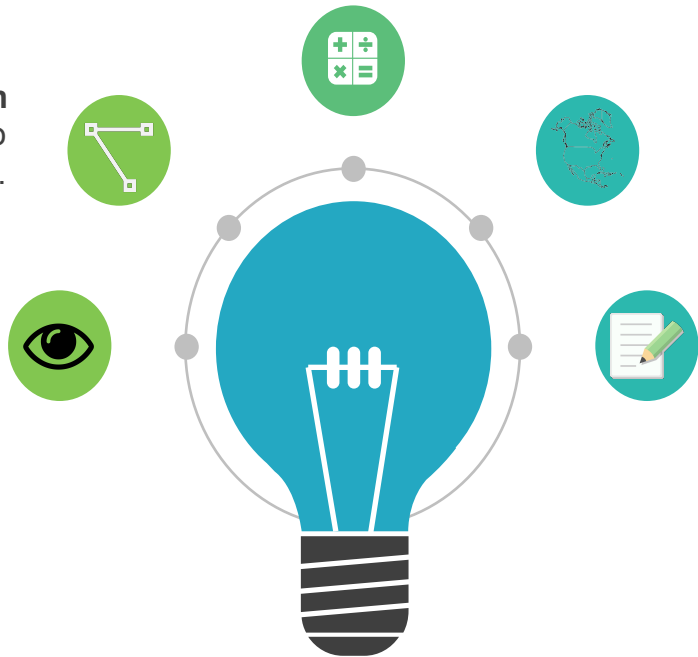
Use a MySQL generic function that calculate shortest distance between two geometries: point-point, point-polyline.

Polyline Simplification

Adapt Douglas-Peucker algorithm to reduce #waypoints by >95%.

Materialized View

Materialize frequently-used join results to accelerate lookups.



Spatial Index

The query optimizer checks the column's SRID attribute and uses appropriate calculations. E.g. R-tree.

Cache & Client-side Filter

Cache intermediate results in client to ensure a smoother user experience.

Optimization

Source	Destination	Distance(mile)	# Accurate Waypoints	# Simplified Waypoints	Time Cost on Google Map	Query Cost before Optimization	Query Cost after Optimization
Levine Hall	MIT	307	5195	13	>10s	>1min	1.6s
Levine Hall	Atlanta, GA	776	8481	19	>10s	>5min	2.2s

Complex Queries

```
SELECT DISTINCT
  sid,
  ST_Y(location) AS longitude,
  ST_X(location) AS latitude
FROM
  materialized_view_electric_stations_denorm E
WHERE
  ST_DISTANCE(
    location,
    ST_SRID(
      ST_GEOFROMTEXT('LINESTRING(
        -75.15062 39.949657,
        -75.142552 39.946464999999996,
        ...
        -75.151775999999998 39.901892000000004
      )',
      4326)
    ) < 804.67
  AND sid IN (SELECT
    sid
    FROM
      station_ports
    WHERE
      port IN ('tesla', 'type2'))
  AND EXISTS(
    SELECT *
    FROM
      electric_stations CLE
    WHERE
      CLE.sid = E.sid
      AND (CLE.ev_level2_evse_num > 0)
  );
```

Built-In Distance Calculator

Simplified Polyline

Find all stations that:

(1) Are close to some path;

(2) Contain some ports;

(3) Support some charging levels.

Complex Queries

CTE

```
WITH S AS (  
    SELECT state FROM stations AS S  
),  
elec_stations AS (  
    SELECT state, count(sid) AS num_stations, 'electric' AS stype FROM  
    ( SELECT DISTINCT S.sid, state FROM electric_stations  
      JOIN stations S on S.sid = electric_stations.sid ) AS A  
    GROUP BY state ORDER BY num_stations DESC  
),  
e85_stations AS (  
    SELECT state, count(sid) AS num_stations, 'e85' AS stype FROM  
    ( SELECT DISTINCT S.sid, state FROM e85  
      JOIN stations S on S.sid = e85.sid ) AS A  
    GROUP BY state ORDER BY num_stations DESC  
# ...  
SELECT * FROM elec_stations  
UNION  
SELECT * FROM e85_stations  
# ...
```

JOIN/GROUP BY

JOIN/GROUP BY

UNION

CTE, multiple aggregation and union over 9 entities.

```
WITH elecStations AS (  
    SELECT sid, state FROM electric_stations  
    NATURAL JOIN stations  
),  
elecStationsPort AS (  
    SELECT * FROM elecStations  
    NATURAL JOIN station_ports  
),  
aggTable AS (  
    SELECT state, COUNT(DISTINCT(sid)) AS numStations, port AS portType  
    FROM elecStationsPort  
    GROUP BY state, portType  
    HAVING state IN ('CA', 'MA') AND portType IN ('type1', '$type2')  
),  
fullTable AS (  
    SELECT * FROM  
    (  
        SELECT 'CA' AS state  
        UNION SELECT 'MA' AS state ) AS stateList,  
    (  
        SELECT 'type1' AS portType  
        UNION SELECT 'type2' AS portType ) AS portList  
)  
SELECT F.state, F.portType, IFNULL(numStations, 0) AS numStations  
FROM fullTable AS F  
LEFT JOIN aggTable AS A  
ON A.state=F.state AND A.portType=F.portType  
ORDER BY state ASC, portType ASC
```

NATURAL JOIN

NATURAL JOIN

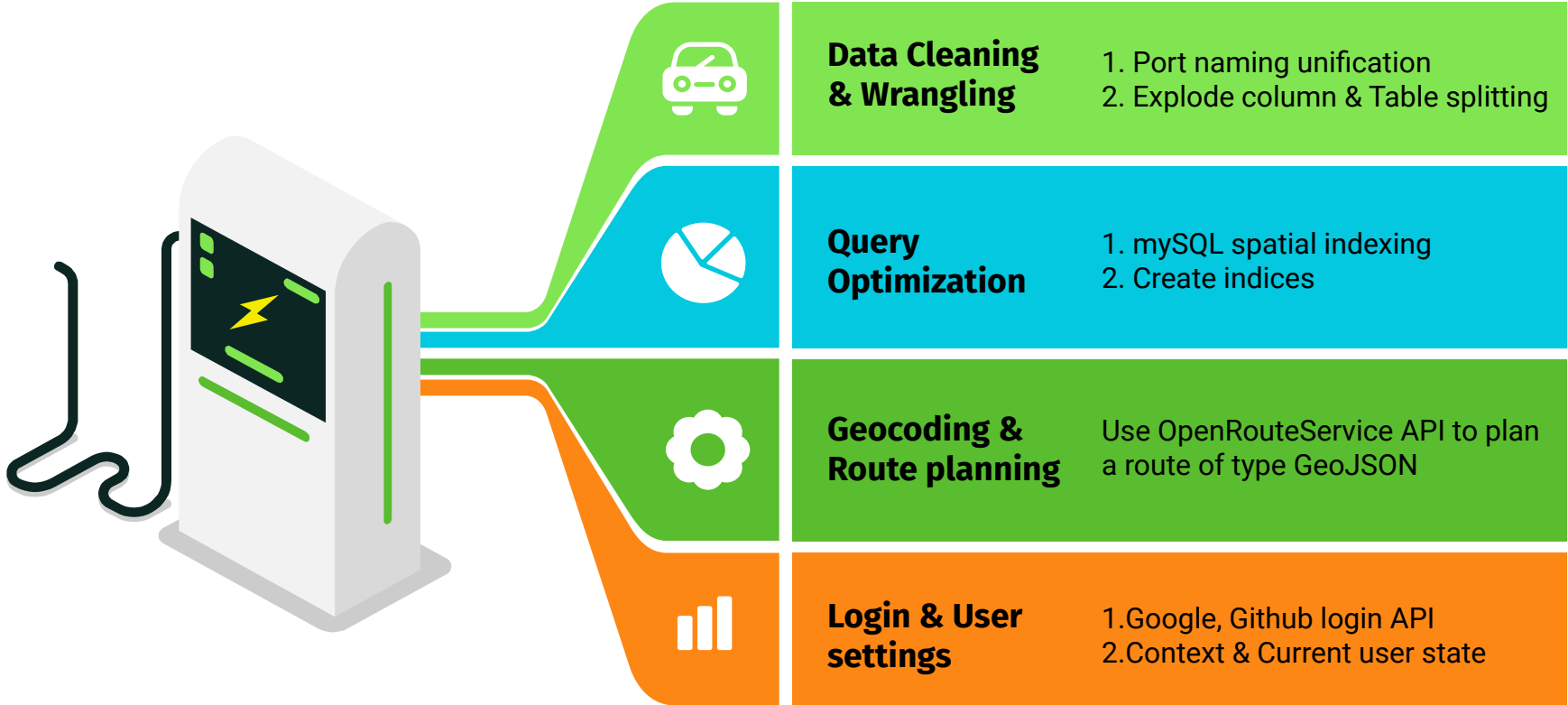
GROUP BY

UNION

LEFT JOIN

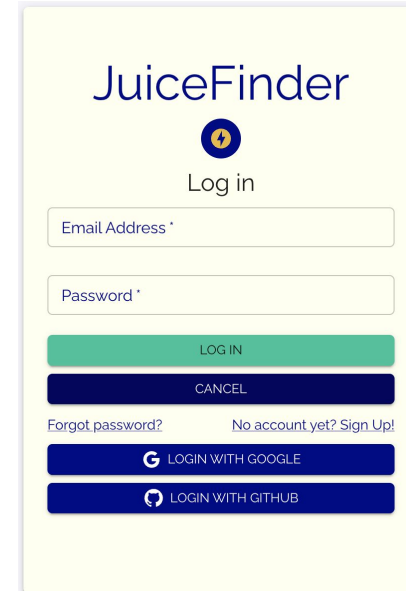
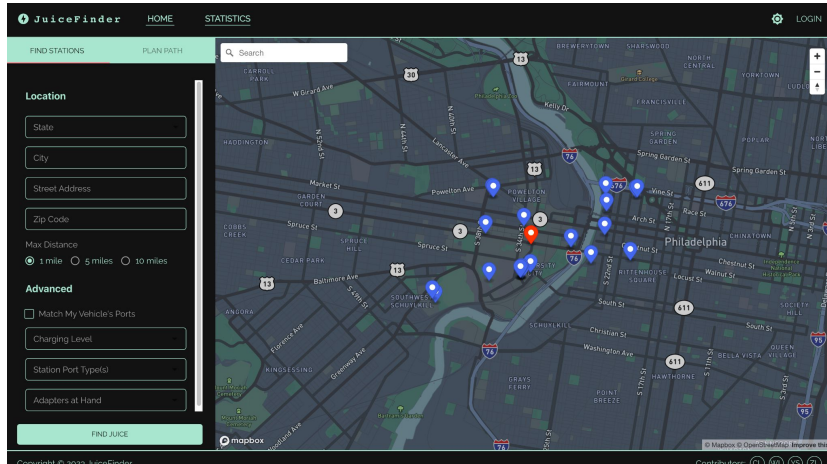
CTE, filtering, aggregation, union, and multiple joins

Technical Challenges



Extra Credits Features

1. Log-in authentication with email/Google/GitHub
2. Light/dark theme switches
3. Extensive visualizations



Thank you!

Q&A